

Serial Number: 09/844,864**ENTERED**CRF Processing Date: 5/11/2001Edited by: AWVerified by: AW

(STIC staff)

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically: _____
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other _____
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: _____
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: _____
- ☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: _____
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: _____
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: _____
- ☒ Deleted: ☒ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as _____
- ☐ Inserted mandatory headings, specifically: _____
- ☐ Corrected an obvious error in the response, specifically: _____
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: _____
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: _____
- ☐ Other: _____

*Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/844,864

DATE: 05/23/2001

TIME: 13:11:58

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05232001\I844864.raw

A 2

2 <110> APPLICANT: Matzuk, Martin
 4 Ren, Yongsheng
 6 Wu, Xuemei
 10 <120> TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
 14 <130> FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
 C--> 18 <140> CURRENT APPLICATION NUMBER: US/09/844,864
 20 <141> CURRENT FILING DATE: 2001-04-27
 24 <150> PRIOR APPLICATION NUMBER: 60/106,020
 26 <151> PRIOR FILING DATE: 1998-10-28
 30 <150> PRIOR APPLICATION NUMBER: PCT/US99/25209
 32 <151> PRIOR FILING DATE: 1999-10-28
 36 <160> NUMBER OF SEQ ID NOS: 25
 40 <170> SOFTWARE: PatentIn version 3.0
 44 <210> SEQ ID NO: 1
 46 <211> LENGTH: 1277
 48 <212> TYPE: DNA
 50 <213> ORGANISM: Mus musculus
 54 <400> SEQUENCE: 1

55	aaggcggg	agggcg	cgaccc	atg	ttccc	ggcg	gcacgt	tcca	cccctg	ccc	60
57	catcctt	atc	cgaggc	ccac	caaagc	cg	gatgg	ctgga	ggttcg	gagc	120
59	cgacccg	cg	ccccct	cct	cccc	cg	tacag	acagc	tcatgg	ccgc	180
61	gacagcc	acc	agcggg	caca	gctcat	ggcc	ctgctg	tcgc	ggatgg	gtcc	240
63	agcagcc	gtg	acgtg	cgggt	gcaggt	gaac	ccgcgc	cg	acgcct	cgggt	300
65	ctcggg	cgcc	gcacgt	tgca	gcctgc	aggg	tgccg	agcca	gccccg	acgc	360
67	tcctgt	caac	cccgtg	ggcca	cgccgg	cgcc	gggag	atccc	cgcgat	cctg	420
69	gccccg	ttct	cgtccg	tgac	cttctg	tggc	ctctc	ctcct	cactgg	aggt	480
71	aggcag	acac	ccacga	aggg	agaggg	gagc	ccggc	atcct	cgggg	acccg	540
73	ccgagag	aggg	tggccg	cgag	gaaagc	ggtc	ccccag	ccgc	gaagc	gagga	600
75	caggctg	cag	ggcagg	ccgg	gtggg	agcag	cagcc	accac	cggagg	accg	660
77	gcggcg	atgc	agtctg	agcc	tgggag	cgag	gagcc	atgtc	ctgccg	caga	720
79	gaccccg	gtg	attcgg	atgc	ccctcg	agac	caggc	ctccc	cgcaa	agcac	780
81	aaggag	cgcc	tgcgtt	tcca	gttctt	tagag	cagaag	tacg	gctact	atca	840
83	tgcaaa	aatcc	ggtggg	agag	cgctat	gtg	tggtgt	gtgc	agggc	accag	900
85	cttcaa	acag	ttctg	ccgag	tgtgtg	agaa	atccta	caac	ccttac	agag	960
87	cacctg	tcaa	agttgt	aaaa	gaacta	gatg	tgcc	tgcca	gtcaga	tttc	1020
89	ccctaa	acgc	ccccat	cggc	aagact	tgtg	tggg	agatgc	aaggac	aaac	1080
91	cgacag	cacc	ttcagc	ttca	aatacat	cat	ttagt	gagag	tcgaaa	acgt	1140
93	tggggc	ttaat	ggaatg	gaca	agtgag	cttt	ctcccc	ctctt	cacctc	ttcc	1200
95	tcttca	tgac	agacag	tggt	acttgg	atat	aaagc	cctgtg	aataaa	aggt	1260
97	aaaaaaaa	aaaaaa							attg	caaaca	1277

100 <210> SEQ ID NO: 2
 102 <211> LENGTH: 361
 104 <212> TYPE: PRT
 106 <213> ORGANISM: Mus musculus
 110 <400> SEQUENCE: 2

112	Met	Phe	Pro	Ala	Ser	Thr	Phe	His	Pro	Cys	Pro	His	Pro	Tyr	Pro	Gln
113	1				5				10					15		

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/844,864

DATE: 05/23/2001

TIME: 13:11:58

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05232001\I844864.raw

```

115 Ala Thr Lys Ala Gly Asp Gly Trp Arg Phe Gly Ala Arg Gly Cys Arg
116                20                25                30
118 Pro Ala Pro Pro Ser Phe Leu Pro Gly Tyr Arg Gln Leu Met Ala Ala
119                35                40                45
121 Glu Tyr Val Asp Ser His Gln Arg Ala Gln Leu Met Ala Leu Leu Ser
122                50                55                60
124 Arg Met Gly Pro Arg Ser Val Ser Ser Arg Asp Ala Ala Val Gln Val
125 65                70                75                80
127 Asn Pro Arg Arg Asp Ala Ser Val Gln Cys Ser Leu Gly Arg Arg Thr
128                85                90                95
130 Leu Gln Pro Ala Gly Cys Arg Ala Ser Pro Asp Ala Arg Ser Gly Ser
131                100               105               110
133 Cys Gln Pro Arg Gly His Ala Gly Ala Gly Arg Ser Pro Arg Ser Trp
134                115               120               125
136 Gln Thr Val Ala Pro Phe Ser Ser Val Thr Phe Cys Gly Leu Ser Ser
137                130               135               140
139 Ser Leu Glu Val Ala Gly Gly Arg Gln Thr Pro Thr Lys Gly Glu Gly
140 145               150               155               160
142 Ser Pro Ala Ser Ser Gly Thr Arg Glu Pro Glu Pro Arg Glu Val Ala
143                165               170               175
145 Ala Arg Lys Ala Val Pro Gln Pro Arg Ser Glu Glu Gly Asp Val Gln
146                180               185               190
148 Ala Ala Gly Gln Ala Gly Trp Glu Gln Gln Pro Pro Pro Glu Asp Arg
149                195               200               205
151 Asn Ser Val Ala Ala Met Gln Ser Glu Pro Gly Ser Glu Glu Pro Cys
152                210               215               220
154 Pro Ala Ala Glu Met Ala Gln Asp Pro Gly Asp Ser Asp Ala Pro Arg
155 225               230               235               240
157 Asp Gln Ala Ser Pro Gln Ser Thr Glu Gln Asp Lys Glu Arg Leu Arg
158                245               250               255
160 Phe Gln Phe Leu Glu Gln Lys Tyr Gly Tyr Tyr His Cys Lys Asp Cys
161                260               265               270
163 Lys Ile Arg Trp Glu Ser Ala Tyr Val Trp Cys Val Gln Gly Thr Ser
164                275               280               285
166 Lys Val Tyr Phe Lys Gln Phe Cys Arg Val Cys Glu Lys Ser Tyr Asn
167                290               295               300
169 Pro Tyr Arg Val Glu Asp Ile Thr Cys Gln Ser Cys Lys Arg Thr Arg
170 305               310               315               320
172 Cys Ala Cys Pro Val Arg Phe Arg His Val Asp Pro Lys Arg Pro His
173                325               330               335
175 Arg Gln Asp Leu Cys Gly Arg Cys Lys Asp Lys Arg Leu Ser Cys Asp
176                340               345               350
178 Ser Thr Phe Ser Phe Lys Tyr Ile Ile
179                355               360
181 <210> SEQ ID NO: 3
183 <211> LENGTH: 1817
185 <212> TYPE: DNA
187 <213> ORGANISM: Mus musculus
191 <400> SEQUENCE: 3

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/844,864

DATE: 05/23/2001
TIME: 13:11:58

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05232001\I844864.raw

```

192 gtcacagctt tcccctgccc gaatatgggtg atctgtctcc attgtccaga tcaggatgat      60
194 tctttagaag aagtcacaga ggaatgctat tccccaccca cctccagaa cctggcaatt      120
196 cagagtctac tgagggatga ggcccttgcc atttctgctc tcacggacct gccccagagt      180
198 ctgttcccag taatttttga ggaggccttc actgatggat atatagggat cttgaaggcc      240
200 atgatacctg tgtggccctt cccatacctt tctttaggaa agcagataaa taattgcaac      300
202 ctggagactt tgaaggctat gcttgaggga ctagatatac tgcttgcaaa aaaggttcaa      360
204 accagtaggt gcaaaactcag agtaattaat tggagagaag atgacttgaa gatatgggct      420
206 ggatcccatg aaggtgaagg cttaccagat ttcaggacag agaagcagcc aattgagaac      480
208 agtgctggct gtgagggtgaa gaaagaattg aaggtgacga ctgaagtcct tcgcatgaag      540
210 ggcagacttg atgaatctac cacatacttg ttgcagtggg cccagcagag aaaagattct      600
212 attcatctat tctgtagaaa gctactaatt gaaggcttaa ccaaagcctc agtgatagaa      660
214 atcttcaaaa ctgtacacgc agactgtata caggagctta tcctaagatg tatctgcata      720
216 gaagagttgg cttttcttaa tccctacctg aaactgatga aaagtctttt cacactcaca      780
218 ctagatcaca tcataggtac cttcagtttg ggtgattctg aaaagcttga tgaggagaca      840
220 atattcagct tgatttctca acttcccaca ctccactgtc tccagaaact ctatgtaaat      900
222 gatgtccctt ttataaaagg caacctgaaa gaatacctca ggtgcctgaa aaagcccttg      960
224 gagacacttt gcatcagtaa ctgtgacctc tcacagtcag acttggattg cctgccctat     1020
226 tgcctgaata tttgtgaact caaacatctg catattagt atatatatt atgtgattta     1080
228 ctcccttgagc ctcttggttt tctccttgag agagttggag ataccctgaa aaccctggaa     1140
230 ttggattcat gttgtatagt ggactttcag ttcagtgcct tgctgcctgc cctaagccaa     1200
232 tgttctcacc tcagagaggt cactttctat gataatgatg tttctctgcc tttcttgaaa     1260
234 acaacttcta caccacacag ccctgctgag tcagctgac tcagctgatt accctgcccc     1320
236 tctagagtgc tatgatgaca gtggtgtaat actaacacac agattagaaa gtttttgtcc     1380
238 tgagcttctg gatatactga gagccaaaag acagctccat agtgtctcct ttcaaacaac     1440
240 caaatgctct aaatgtggtg ggtgctacat ttatgatcgg catacccaat gttgccgttt     1500
242 tgtggaacta ctataagctt gattgtgaaa ctgagaaata gaaacttagt attggggact     1560
244 gatgaaatcc taagtgaatg tccactgcta aatggagcat gaaaatgtca atcacctaaa     1620
246 agtctgagat acacaggaaa gtcaataact tcctctgagc tgggtgaatg atgttgcac     1680
248 tgtagaaagt atcaagcact tgtagtttga atgtgttaca atagaagcac cattttatga     1740
250 gactggccca atctgttgac tgcatacaat aaatctgttg acttattaaa tttttaaaaa     1800
252 aaaaaaaaaa aaaaaaa
255 <210> SEQ ID NO: 4
257 <211> LENGTH: 426
259 <212> TYPE: PRT
261 <213> ORGANISM: Mus musculus
265 <400> SEQUENCE: 4
267 Met Val Ile Cys Leu His Cys Pro Asp Gln Asp Asp Ser Leu Glu Glu
268 1 5 10 15
270 Val Thr Glu Glu Cys Tyr Ser Pro Pro Thr Leu Gln Asn Leu Ala Ile
271 20 25 30
273 Gln Ser Leu Leu Arg Asp Glu Ala Leu Ala Ile Ser Ala Leu Thr Asp
274 35 40 45
276 Leu Pro Gln Ser Leu Phe Pro Val Ile Phe Glu Glu Ala Phe Thr Asp
277 50 55 60
279 Gly Tyr Ile Gly Ile Leu Lys Ala Met Ile Pro Val Trp Pro Phe Pro
280 65 70 75 80
282 Tyr Leu Ser Leu Gly Lys Gln Ile Asn Asn Cys Asn Leu Glu Thr Leu
283 85 90 95
285 Lys Ala Met Leu Glu Gly Leu Asp Ile Leu Leu Ala Gln Lys Val Gln

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/844,864

DATE: 05/23/2001

TIME: 13:11:58

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05232001\I844864.raw

```

286          100          105          110
288 Thr Ser Arg Cys Lys Leu Arg Val Ile Asn Trp Arg Glu Asp Asp Leu
289          115          120          125
291 Lys Ile Trp Ala Gly Ser His Glu Gly Glu Gly Leu Pro Asp Phe Arg
292          130          135          140
294 Thr Glu Lys Gln Pro Ile Glu Asn Ser Ala Gly Cys Glu Val Lys Lys
295 145          150          155          160
297 Glu Leu Lys Val Thr Thr Glu Val Leu Arg Met Lys Gly Arg Leu Asp
298          165          170          175
300 Glu Ser Thr Thr Tyr Leu Leu Gln Trp Ala Gln Gln Arg Lys Asp Ser
301          180          185          190
303 Ile His Leu Phe Cys Arg Lys Leu Leu Ile Glu Gly Leu Thr Lys Ala
304          195          200          205
306 Ser Val Ile Glu Ile Phe Lys Thr Val His Ala Asp Cys Ile Gln Glu
307          210          215          220
309 Leu Ile Leu Arg Cys Ile Cys Ile Glu Glu Leu Ala Phe Leu Asn Pro
310 225          230          235          240
312 Tyr Leu Lys Leu Met Lys Ser Leu Phe Thr Leu Thr Leu Asp His Ile
313          245          250          255
315 Ile Gly Thr Phe Ser Leu Gly Asp Ser Glu Lys Leu Asp Glu Glu Thr
316          260          265          270
318 Ile Phe Ser Leu Ile Ser Gln Leu Pro Thr Leu His Cys Leu Gln Lys
319          275          280          285
321 Leu Tyr Val Asn Asp Val Pro Phe Ile Lys Gly Asn Leu Lys Glu Tyr
322          290          295          300
324 Leu Arg Cys Leu Lys Lys Pro Leu Glu Thr Leu Cys Ile Ser Asn Cys
325 305          310          315          320
327 Asp Leu Ser Gln Ser Asp Leu Asp Cys Leu Pro Tyr Cys Leu Asn Ile
328          325          330          335
330 Cys Glu Leu Lys His Leu His Ile Ser Asp Ile Tyr Leu Cys Asp Leu
331          340          345          350
333 Leu Leu Glu Pro Leu Gly Phe Leu Leu Glu Arg Val Gly Asp Thr Leu
334          355          360          365
336 Lys Thr Leu Glu Leu Asp Ser Cys Cys Ile Val Asp Phe Gln Phe Ser
337          370          375          380
339 Ala Leu Leu Pro Ala Leu Ser Gln Cys Ser His Leu Arg Glu Val Thr
340 385          390          395          400
342 Phe Tyr Asp Asn Asp Val Ser Leu Pro Phe Leu Lys Thr Thr Ser Thr
343          405          410          415
345 Pro His Ser Pro Ala Glu Ser Ala Asp Leu
346          420          425
348 <210> SEQ ID NO: 5
350 <211> LENGTH: 1018
352 <212> TYPE: DNA
354 <213> ORGANISM: Mus musculus
358 <400> SEQUENCE: 5
359 gccatattga ggacctgcag tagaggtgga acccatgact ggcagcgcaa acacagtgat      60
361 aacagctgag ctccaagcaa ggaccagga ccttgctca ccacagacat aatctttccc      120
363 cacaacacct ccaccaagcc gccctgtaaa tcgacatgag tcgccacagc accagcagcg      180

```


RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/844,864

DATE: 05/23/2001
TIME: 13:11:58

Input Set : A:\Pto.amc
Output Set: C:\CRF3\05232001\I844864.raw

```

365 tgaccgaaac cacagcaaaa aacatgctct ggggtagtga actcaatcag gaaaagcaga 240
367 cttgcacctt tagaggccaa ggcgagaaga aggacagctg taaactcttg ctcagcacga 300
369 tctgcctggg ggagaaagcc aaagaggagg tgaaccgtgt ggaagtcctc tcccaggaag 360
371 gcagaaaacc accaatcact attgctacgc tgaaggcatc agtcctgccc atggtcactg 420
373 tgtcaggtat agagctttct cctccagtaa cttttcggct caggactggc tcaggacctg 480
375 tgttcctcag tggcctggaa tgttatgaga cttcggacct gacctgggaa gatgacgagg 540
377 aagaggagga agaggaggag gaagaggatg aagatgagga tgcagatata tcgctagagg 600
379 agatacctgt caaacaagtc aaaaggggtg ctccccagaa gcagatgagc atagcaaaga 660
381 aaaagaagggt ggaaaaagaa gaggatgaaa cagtagtgag gccagccct caggacaaga 720
383 gtccctggaa gaaggagaaa tctacacca gagcaaagaa gccagtgacc aagaaatgac 780
385 ctcactcttag catcttctgc gtccaaggca ggatgtccag cagctgtgtt ttggtgcagg 840
387 tgtccagccc caccacccta gtctgaatgt aataagggtg tgtggctgta accctgtaac 900
389 ccagccctcc agtttccgga ggtttttggt gaagagcccc cagcaagttc gcctagggcc 960
391 acaataaaat ttgcatgata aggaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 1018
394 <210> SEQ ID NO: 6
396 <211> LENGTH: 207
398 <212> TYPE: PRT
400 <213> ORGANISM: Mus musculus
404 <400> SEQUENCE: 6
406 Met Ser Arg His Ser Thr Ser Ser Val Thr Glu Thr Thr Ala Lys Asn
407 1 5 10 15
409 Met Leu Trp Gly Ser Glu Leu Asn Gln Glu Lys Gln Thr Cys Thr Phe
410 20 25 30
412 Arg Gly Gln Gly Glu Lys Lys Asp Ser Cys Lys Leu Leu Leu Ser Thr
413 35 40 45
415 Ile Cys Leu Gly Glu Lys Ala Lys Glu Glu Val Asn Arg Val Glu Val
416 50 55 60
418 Leu Ser Gln Glu Gly Arg Lys Pro Pro Ile Thr Ile Ala Thr Leu Lys
419 65 70 75 80
421 Ala Ser Val Leu Pro Met Val Thr Val Ser Gly Ile Glu Leu Ser Pro
422 85 90 95
424 Pro Val Thr Phe Arg Leu Arg Thr Gly Ser Gly Pro Val Phe Leu Ser
425 100 105 110
427 Gly Leu Glu Cys Tyr Glu Thr Ser Asp Leu Thr Trp Glu Asp Asp Glu
428 115 120 125
430 Glu Glu Glu Glu Glu Glu Glu Glu Asp Glu Asp Glu Asp Ala Asp
431 130 135 140
433 Ile Ser Leu Glu Glu Ile Pro Val Lys Gln Val Lys Arg Val Ala Pro
434 145 150 155 160
436 Gln Lys Gln Met Ser Ile Ala Lys Lys Lys Lys Val Glu Lys Glu Glu
437 165 170 175
439 Asp Glu Thr Val Val Arg Pro Ser Pro Gln Asp Lys Ser Pro Trp Lys
440 180 185 190
442 Lys Glu Lys Ser Thr Pro Arg Ala Lys Lys Pro Val Thr Lys Lys
443 195 200 205
445 <210> SEQ ID NO: 7
447 <211> LENGTH: 214
449 <212> TYPE: DNA
451 <213> ORGANISM: Mus musculus

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/844,864

DATE: 05/23/2001

TIME: 13:11:59

Input Set : A:\Pto.amc

Output Set: C:\CRF3\05232001\I844864.raw

L:18 M:270 C: Current Application Number differs, Replaced Current Application Number